Randomness and Initial Segment Complexity for Probability Measures

André Nies
School of Computer Science, The University of Auckland, New Zealand
https://www.cs.auckland.ac.nz/~nies/
andre@cs.auckland.ac.nz

Frank Stephan
Department of Mathematics and Department of Computer Science,
National University of Singapore, Singapore
fstephan@comp.nus.edu.sg

Abstract

We study algorithmic randomness properties for probability measures on Cantor space. We say that a measure μ on the space of infinite bit sequences is Martin-Löf absolutely continuous if the non-Martin-Löf random bit sequences form a null set with respect to μ. We think of this as a weak randomness notion for measures. We begin with examples, and a robustness property related to Solovay tests. Our main work connects our property to the growth of the initial segment complexity for measures μ; the latter is defined as a μ-average over the complexity of strings of the same length. We show that a maximal growth implies our weak randomness property, but also that both implications of the Levin-Schnorr theorem fail. We briefly discuss K-triviality for measures, which means that the growth of initial segment complexity is as slow as possible. We show that full Martin-Löf randomness of a measure implies Martin-Löf absolute continuity; the converse fails because only the latter property is compatible with having atoms. In a final section we consider weak randomness relative to a general ergodic computable measure. We seek appropriate effective versions of the Shannon-McMillan-Breiman theorem and the Brudno theorem where the bit sequences are replaced by measures.

2012 ACM Subject Classification Theory of computation; Theory of computation → Quantum information theory

Keywords and phrases algorithmic randomness, probability measure on Cantor space, Kolmogorov complexity, statistical superposition, quantum states

Digital Object Identifier 10.4230/LIPIcs.STACS.2020.55

Funding *André Nies*: Partially supported by the Marsden Fund of the Royal Society of New Zealand, UOA 13-184.
Frank Stephan: Partially supported by the Singapore Ministry of Education Academic Research Fund Tier 2 grant MOE2016-T2-1-019 / R146-000-234-112.

1 Introduction

The theory of algorithmic randomness is usually developed for bit sequences. A central randomness notion based on algorithmic tests is the one due to Martin-Löf [11].

Let $\{0,1\}^{\mathbb{N}}$ denote the topological space of infinite bit sequences. A probability measure μ on $\{0,1\}^{\mathbb{N}}$ can be seen as a statistical superposition of bit sequences. The bit sequences Z form an extreme case: the corresponding measure μ is the Dirac measure δ_Z, i.e., μ is concentrated on $\{Z\}$. The opposite extreme is the uniform measure λ which independently gives each bit value the probability $1/2$. The uniform measure represents the maximum disorder as no bit sequence is preferred over any other.

Recall that a measure μ on $\{0,1\}^{\mathbb{N}}$ is called *absolutely continuous* if each λ-null set is a μ-null set. Our main concept is an algorithmic randomness notion for probability measures that is a weakening of absolute continuity: we require that the λ-null set in the hypothesis be effective in the sense of Martin-Löf. Given that there is a universal Martin-Löf test, and hence a largest effective null set, all we have to require is that $\mu(\mathcal{C}) = 0$ where $\mathcal{C}$ is the class of bit sequences that are not Martin-Löf random.

Our research is partly motivated by a recent definition of Martin-Löf randomness for quantum states corresponding to infinitely many qubits, due to the first author and Scholz [18]. Using the terminology there, probability measures correspond to the quantum states ρ where the matrix $\rho\restriction_{M_n}$ is diagonal for each n, where M_n is the algebra of complex $2^n \times 2^n$ matrices. Subsequent work of Tejas Bhojraj has shown that for measures, the randomness notion defined there is equivalent to the one proposed here. So the measures form a useful intermediate case to test conjectures in the subtler setting of quantum states. This applies, for instance, to the SMB theorem discussed at the end of this section, which is studied in the setting of quantum states by the first author and Tomamichel (see the post in [20]).

Growth of initial segment complexity. Given a binary string x, by $C(x)$ one denotes its plain descriptive complexity, and by $K(x)$ its prefix-free descriptive complexity. Our main motivation is derived from the classical theory. Randomness of infinite bit sequences is linked to the growth of the descriptive complexity of their initial segments. For instance, the Levin-Schnorr theorem intuitively says that randomness of Z means incompressibility (up to the same constant b) of all the initial segments of Z. We want to study how much of this is retained in the setting of measures μ. One now takes the μ-average over the complexity of all strings of a given length n. It turns out that interesting new growth behaviour is possible, such as having maximal growth of C-complexity on all initial segments. This growth rate is ruled out for bit sequences by a result of Katseff. However, using that "most strings are incompressible" we verify in Fact 9 that the uniform measure λ has this growth behaviour, namely $C(\lambda\restriction_n) \geq^+ n$. On the other hand, we show that this type of fast growth implies our weak randomness notion.

The formal growth condition in the Levin-Schnorr theorem says that $K(x) \geq |x| - b$ for each initial segment x of Z, where $K(x)$ is the prefix free version of Kolmogorov complexity of a string x. The "n-th initial segment" of a measure μ is given by its values $\mu[x]$ for all strings x of length n, where $[x]$ denotes the set of infinite sequences extending x. As mentioned, it is natural to define the initial segment complexities $C(\mu\restriction_n)$ and $K(\mu\restriction_n)$ of this initial segment as the μ-average of the individual complexities of those strings. With this definition, in Section 3 we show that both implications of the analog of the Levin-Schnorr theorem fail. However, we also show in Proposition 25 that for measures that are random in our weak sense, $C(\mu\restriction_n)/n$, or equivalently $K(\mu\restriction_n)/n$, converges to 1. Thus, such measures have effective dimension 1; see Downey and Hirschfeldt [5, Section 12.3] on effective dimension.

Further results and potential avenues for future research. Opposite to random bit sequences are the K-trivial sequences, where the initial segment complexity grows no faster than that of a computable set; for background see e.g. Nies [17, Section 5.3]. In Section 4 we briefly extend this notion to statistical superpositions of bit sequences: we introduce K-trivial measures and show that they have countable support. This means that they are countable combinations of Dirac measures.

Measures can be viewed points in a canonical computable probability space, in the sense of [8]. This yields a notion of Martin-Löf randomness for measures. Culver [4, Th. 2.7.1] has shown that no measure μ that is Martin-Löf random is absolutely continuous; in fact μ

is orthogonal to λ in the sense that some null set is co-null with respect to μ. In contrast, in Section 5 we show that this notion implies our weak notion of randomness, ML-absolute continuity. The stronger randomness notion forces the measure to be atomless, so the converse implication fails. Further questions can be asked about the relationships between the different randomness notions for measures we have discussed. For instance, does the strong notion imply maximal growth of initial segment complexity for the measure (in the sense of C or of K)? We plan to address such questions in the upcoming journal version of the paper.

The Shannon-McMillan-Breiman Theorem from the 1950s (see [24], where it is called the Entropy Theorem) says informally that for an ergodic measure ρ on $\{0,1\}^{\mathbb{N}}$, outside a null set every bit sequence Z reflects the entropy of the measure ρ by the limiting weighted information content on its sufficiently large initial segments. In the final Section 6 we study what happens when Z is replaced by a measure μ that is Martin-Löf a.c. with respect to ρ and we take the μ-average of the information contents at the same length. Here we only obtain a partial result. However, in a similar vein, we establish an analog for measures of the effective Brudno's theorem [6, 7] that the entropy of ρ is given as the limit of $K(Z\restriction_n)/n$, for any Z that is ρ-ML random. Obtaining a measure version of the effective Birkhoff ergodic theorem would be interesting as well.

For general background on recursion theory and algorithmic randomness we refer the readers to the textbooks of Calude [2], Downey and Hirschfeldt [5], Li and Vitányi [10], Nies [17], Odifreddi [21, 22], and Soare [25]. Lecture notes on recursion theory are also available online, e.g. [26].

2 Measures and Randomness

In this section we formally define our main notion (Definition 3), and collect some basic facts concerning it. In particular, we verify that the well-known equivalence of Martin-Löf test and Solovay tests extends to measures. We begin by briefly discussing algorithmic randomness for bit sequences [5, 17]. We use standard notation: letters $Z, X, \ldots$ denote elements of the space of infinite bit sequences $\{0,1\}^{\mathbb{N}}$, σ, τ denote finite bit strings, and $[\sigma] = \{Z\colon Z \succ \sigma\}$ is the set of infinite bit sequences extending σ. $Z\restriction_n$ denotes the string consisting of the first n bits of Z. For quantities r, s depending on the same parameters, one writes $r \le^+ s$ for $r \le s + O(1)$. A subset G of $\{0,1\}^{\mathbb{N}}$ is called *effectively open* if $G = \bigcup_i [\sigma_i]$ for a computable sequence $\langle \sigma_i \rangle_{i\in\mathbb{N}}$ of strings. A measure ρ on $\{0,1\}^{\mathbb{N}}$ is computable if the map $\{0,1\}^{<\omega} \to \mathbb{R}$ given by $\sigma \mapsto \rho[\sigma]$ is computable.

▶ **Definition 1.** *Let ρ be a computable measure on $\{0,1\}^{\mathbb{N}}$. A ρ-Martin-Löf test (ρ-ML-test, for short) is a sequence $\langle G_m \rangle$ of uniformly effectively open sets such that $\rho G_m \le 2^{-m}$ for each m. A bit sequence Z* fails *the test if $Z \in \bigcap_m G_m$, otherwise it* passes *the test. A bit sequence Z is ρ- Martin-Löf random (ML-random) if Z passes each ρ-ML-test.*

By λ one denotes the uniform measure on $\{0,1\}^{\mathbb{N}}$. So $\lambda[\sigma] = 2^{-|\sigma|}$ for each string σ. If no measure ρ is provided it will be tacitly assumed that $\rho = \lambda$, and we will use the term ML-random instead of λ-ML-random etc. Let $K(x)$ denote the prefix free version of descriptive (i.e., Kolmogorov) complexity of a bit string x.

▶ **Theorem 2** (Levin [9], Schnorr [23]). *Z is ML-random $\Leftrightarrow \exists b \forall n K(Z\restriction_n) \ge n - b$.*

Using the notation of [17, Ch. 3], let $\mathcal{R}_b$ denote the set of bit sequences Z such that $K(Z\restriction_n) < n - b$ for some n. It is easy to see that $\langle \mathcal{R}_b \rangle_{b\in\mathbb{N}}$ forms a Martin-Löf test. The Levin-Schnorr theorem says that this test is universal: Z is ML-random iff it passes the test.

Unless otherwise stated, all measures will be probability measures. We use the letters μ, ν, ρ for probability measures (and recall that λ denotes the uniform measure). We now provide the formal definition of our weak randomness notion for measures.

▶ **Definition 3** (Main). *A measure μ is called* Martin-Löf absolutely continuous *in ρ (ρ-ML a.c., for short) if $\inf_m \mu(G_m) = 0$ for each ρ-Martin-Löf test $\langle G_m \rangle_{m \in \mathbb{N}}$. We write $\mu \ll_{ML} \rho$.*

If $\inf_m \mu(G_m) = 0$ we say that μ *passes* the test. If $\inf_m \mu(G_m) \geq \delta$ where $\delta > 0$ we say μ *fails the test at level δ*.

Martin-Löf absolute continuity is a *weakening* of the usual notion of absolute continuity $\mu \ll \rho$. In fact, $\mu \ll \rho$ iff μ is ρ-ML^X-a.c. for each oracle X.

In the definition it suffices to consider ρ-ML tests $\langle G_m \rangle$ such that $G_m \supseteq G_{m+1}$ for each m, because we can replace $\langle G_m \rangle$ by the ρ-ML test $\widehat{G}_m = \bigcup_{k>m} G_k$, and of course $\inf_m \mu(\widehat{G}_m) = 0$ implies $\inf_m \mu(G_m) = 0$. So we can change the definition above, replacing the condition $\inf_m G_m = 0$ by the only apparently stronger condition $\lim_m G_m = 0$.

The intersection of a universal ρ-ML test consists of the non-ML random sequences. Since such a test exists, we have:

▶ **Fact 4.** *$\mu \ll_{ML} \rho$ iff the sequences which are not ρ-ML random form a μ-null set.*

We have already mentioned the two diametrically opposite types of examples:

▶ **Example 5.** (a) $\rho \ll_{ML} \rho$.

(b) For a Dirac measure δ_Z, we have $\delta_Z \ll_{ML} \rho$ iff Z is ρ-ML random.

For $p \neq 1/2$, a Bernoulli measure on $\{0,1\}^{\mathbb{N}}$, that independently gives probability p to a 0 in each position, is not Martin-Löf a.c. To see this, note that each ML-random sequence Z satisfies the law of large numbers

$$\lim_n \frac{1}{n} |\{i < n \colon Z(i) = 1\}| = 1/2;$$

see e.g. [17, Prop. 3.2.19]. So if μ is Martin-Löf a.c., then μ-almost surely, Z satisfies the law of large numbers. This is not the case for such Bernoulli measures.

For a measure ν and string σ with $\nu[\sigma] > 0$ let ν_σ be the localisation:

$$\nu_\sigma(A) = \nu(A \cap [\sigma]) / \nu[\sigma].$$

Clearly if ν is Martin-Löf a.c. then so is ν_σ.

A set S of probability measures is called *convex* if $\mu_i \in S$ for $i \leq k$ implies that the convex combination $\mu = \sum_i \alpha_i \mu_i$ is in S, where the α_i are reals in $[0,1]$ and $\sum_i \alpha_i = 1$. The extreme points of S are the ones that can only be written as convex combinations of length 1 of elements of S.

▶ **Proposition 6.** *The Martin-Löf a.c. probability measures form a convex set. Its extreme points are the Martin-Löf a.c. Dirac measures, i.e. the measures δ_Z where Z is a ML-random bit sequence.*

Proof. Let $\mu = \sum_i \alpha_i \mu_i$ as above where the μ_i are Martin-Löf a.c. measures. Suppose $\langle G_m \rangle$ is a Martin-Löf test. Then $\lim_m \mu_i(G_m) = 0$ for each i, and hence $\lim_m \mu(G_m) = 0$.

Suppose that μ is Martin-Löf a.c. If μ is a Dirac measure then it is an extreme point of the Martin-Löf a.c. measures. Conversely, if μ is not Dirac, there is a least number t such that the decomposition

$$\mu = \sum_{|\sigma| = t, \mu[\sigma] > 0} \mu[\sigma] \cdot \mu_\sigma$$

is nontrivial. Hence μ is not an extreme point. ◀

Recall that a *Solovay test* is a sequence $\langle S_k \rangle_{k\in\mathbb{N}}$ of uniformly Σ^0_1 sets such that $\sum_k \lambda(S_k) < \infty$. A bit sequence Z passes such a test if $Z \notin S_k$ for almost every k. (Each ML-test is a Solovay test, but the passing condition is stronger for Solovay tests). A basic fact from the theory of algorithmic randomness (e.g. [17, 3.2.9]) states that Z is ML-random iff Z passes each Solovay test.

The following characterises the Martin-Löf a.c. measures with countable support.

▶ **Fact 7.** *Let $\mu = \sum_k c_k \delta_{Z_k}$ where $\forall k\, [0 < c_k \le 1]$ and $\sum_k c_k = 1$. Then μ is Martin-Löf a.c. iff all the Z_k are Martin-Löf random.*

Proof. The implication from left to right is immediate. For the converse implication, given a Martin-Löf test $\langle G_m \rangle$, note that the Z_k pass this test as a Solovay test. Hence for each r, there is M such that $Z_k \notin G_m$ for each $k \le r$ and each $m \ge M$. This implies that $\mu(G_m) \le \sum_{k>r} c_k$ for each $m \ge M$. So $\lim_m \mu(G_m) = 0$. ◀

We say that a measure μ *passes* a Solovay test $\langle S_k \rangle_{k\in\mathbb{N}}$ if $\lim_k \mu(S_k) = 0$. The fact that passing all Martin-Löf tests is equivalent to passing all Solovay tests generalises from bit sequences to measures. We note that Tejas Bhojraj (in preparation) proved such a result in even greater generality in the setting of quantum states, where the proof is more involved.

▶ **Proposition 8.** *A measure μ is Martin-Löf a.c. iff μ passes each Solovay test.*

Proof. Each Martin-Löf test is a Solovay test, and the passing condition $\lim_m \mu(G_m) = 0$ works for both types of tests by the remark after Definition 3. This yields the implication from right to left.

For the implication from left to right, let μ be Martin-Löf a.c. and let $\langle S_k \rangle_{k\in\mathbb{N}}$ be a Solovay test. By $\limsup_k S_k$ one denotes the set of bit sequences Z such that $\exists^\infty k\, Z \in S_k$, that is, the sequences that fail the test. By the basic fact (e.g. [17, 3.2.9]) mentioned above, the set of ML-random sequences is disjoint from $\limsup_k S_k$. By hypothesis on μ we have $\mu(\limsup_k S_k) = 0$. By Fatou's Lemma, $\limsup_k \mu(S_k) \le \mu(\limsup_k S_k)$. So μ passes the Solovay test. ◀

3 Initial segment complexity of a measure μ

Let $K(\mu \upharpoonright_n) = \sum_{|x|=n} K(x)\mu[x]$ be the μ-average of all the $K(x)$ over all strings x of length n. In a similar way we define $C(\mu \upharpoonright_n)$. Note that for a Dirac measure δ_Z, we have $K(\delta_Z \upharpoonright_n) = K(Z \upharpoonright_n)$.

In this section we use standard inequalities such as $C(x) \le^+ K(x)$, $K(x) \le^+ |x| + 2\log|x|$ and $K(0^n) \le^+ 2\log n$. We also use that for each r there are at most $2^r - 1$ strings such that $C(x) < r$. See e.g. [17, Ch. 2]. Recall that λ denotes the uniform measure on $\{0,1\}^{\mathbb{N}}$.

3.1 A fast growing initial segment complexity implies being ML-a.c.

Recall that $C(x) \le^+ |x|$ and $K(x) \le^+ |x| + K(|x|)$. The following says that the uniform measure λ has the fastest growing initial segment complexity that is possible in both sense of C and of K.

▶ **Fact 9.**
(a) $C(\lambda \upharpoonright_n) \ge^+ n$.
(b) $K(\lambda \upharpoonright_n) \ge^+ n + K(n)$.

Proof. Chaitin [3] showed that there is a constant c such that, for all d, there are at most 2^{n+c-d} strings $x \in \{0,1\}^n$ with $C(x) \le n-d$. Similarly, among the strings of length n, there are at most 2^{n+c-d} strings with $K(x) \le n+K(n)-d$. In other words, the fraction of strings of length n where, for (a), $C(x) \le n-d$, and, for (b), $K(x) \le n+K(n)-d$, respectively, is in each case at most 2^{c-d}. Now for each d, from the estimated lower bound n and $n+K(n)$, respectively, one subtracts the fraction of the strings of length n for which the Kolmogorov complexity is at least d below the average in order to correct the lower bound. For, if x is a string of length n such that $C(x) \le n-r$ (resp, $K(x) \le n+K(n)-r$), then in computing $C(\lambda\restriction_n)$ (resp, $K(\lambda\restriction_n)$) a correction of 2^{-n} has to be subtracted r times, for $d = 1,\dots,r$.

Let c_d be the fraction of strings of length n with $C(x) \le n-d$, and let k_d be the fraction of strings with $K(x) \le n+K(n)-d$. Then as argued above,

$$C(\lambda\restriction_n) \ge n - \sum_{d\ge 0} c_d \text{ and } K(\lambda\restriction_n) \ge n + K(n) - \sum_{d\ge 0} k_d.$$

Using Chaitin's bounds gives then the corrected estimates on the averages of

$$C(\lambda\restriction_n) \ge n - \sum_{d\ge 0} 2^{c-d} \text{ and } K(\lambda\restriction_n) \ge n + K(n) - \sum_{d\ge 0} 2^{c-d}.$$

Now one uses that $\sum_{d\ge 0} 2^{c-d} \le 2^{c+1}$ and that 2^{c+1} is a constant independent of n and only dependent on the universal machine in order to get that $C(\lambda\restriction_n) \ge^+ n$ and $K(\lambda\restriction_n) \ge^+ n+K(n)$. ◀

Recall [10] that a bit sequence $Z \in \{0,1\}^{\mathbb{N}}$ is called Kolmogorov random if there is r such that $C(Z\restriction_n) \ge n-r$ for infinitely many n; Z is strongly Chaitin random if there is r such that $K(Z\restriction_n) \ge n+K(n)-r$ for infinitely many n. For bit sequences these notions are equivalent to 2-randomness by [19] and [13], respectively; also see [17, 8.1.14] or [5].

We may extend these notion to measures. We show that a measure satisfying either of the notions is Martin-Löf a.c.:

▶ **Theorem 10.**
(a) *Suppose that μ is a measure such that there is an r with $C(\mu\restriction_n) \ge n-r$ for infinitely many n. Then μ is Martin-Löf a.c.*
(b) *The same conclusion holds under the hypothesis that $K(\mu\restriction_n) \ge n+K(n)-r$ for infinitely many n.*

Proof. Suppose that μ is not Martin-Löf a.c. So there is a Martin-Löf test $\langle G_d\rangle_{d\in\mathbb{N}}$ and $\delta > 0$ such that $\mu(G_d) \ge \delta$ for each d. We view G_d as given by an enumeration of strings, uniformly in d; thus $G_d = \bigcup_i [\sigma_i]$ for a sequence $\langle \sigma_i\rangle_{i\in\mathbb{N}}$ that is computable uniformly in d. Let $G_d^{\le n}$ denote the clopen set generated by the strings in this enumeration of length at most n. (Note that this set is not effectively given as a clopen set, but we effectively have a description of it as a Σ^0_1 set). Let c be a constant such that, for each x of length n, one has $C(x) \le n+c$ in case (a), and $K(x) \le n+K(n)+c$ in case (b).

▶ **Lemma 11.** *If x is a string of length n such that $[x] \subseteq G_d^{\le n}$ then $C(x \mid d) \le^+ n-d$ and $K(x \mid n,d) \le^+ n-d$.*

To verify this, we first consider the case of plain complexity C. Let N be a fixed plain machine that on input y and auxiliary input d prints out the y-th string x of length $n = |y|+d$ such that our enumeration of $G_d^{\le n}$ asserts that $[x] \subseteq G_d^{\le n}$. (Here we view y as the binary representation of a number, with leading zeros allowed.) Since $\lambda G_d \le 2^{-d}$, sufficiently many strings are available to print all such x. This machine shows that $C(x \mid d) \le^+ n-d$ for any x such that $[x] \subseteq G_d^{\le n}$, as required.

For the case of prefix free complexity K, let N' be the slightly modified machine where both n and d are auxiliary inputs. The machine N provides for a string x of length n a description of length $n-d$. So for N', for the same pair n, d, the descriptions of different strings form a prefix free set. This verifies the lemma.

Now such x satisfy (after increaing x, if necessary) that $C(x) \le n - d + 2\log d + c$ and $K(x) \le n + K(n) - d + 2\log d + c$. We complete the proof separately for (a) and (b).

(a) For each d, n, letting x range over strings of length n, we have

$$C(\mu\restriction_n) = \sum_{|x|=n} C(x)\mu[x] = \sum_{[x]\subseteq G_d^{\le n}} C(x)\mu[x] + \sum_{[x]\not\subseteq G_d^{\le n}} C(x)\mu[x].$$

The first summand is bounded above by $\mu(G_d^{\le n})(n - d + 2\log d + c)$ via the lemma, the second by $(1-\mu(G_d^{\le n}))(n+c)$. We obtain

$$C(\mu\restriction_n) \le n + c - \mu(G_d^{\le n})d/2.$$

Now for each d, for sufficiently large n we have $\mu(G_d^{\le n}) \ge \delta$. So given r let $d = 2r/\delta$; then for large enough n we have $C(\mu\restriction_n) \le n + c - r$.

(b) For each d, n, letting x range over strings of length n, we have

$$K(\mu\restriction_n) = \sum_{|x|=n} K(x)\mu[x] = \sum_{[x]\subseteq G_d^{\le n}} K(x)\mu[x] + \sum_{[x]\not\subseteq G_d^{\le n}} K(x)\mu[x].$$

The first summand is bounded above by $\mu(G_d^{\le n})(n + K(n) - d + 2\log d + c)$, the second by $(1-\mu(G_d^{\le n}))(n + K(n) + c)$. We obtain

$$K(\mu\restriction_n) \le n + K(n) + c - \mu(G_d^{\le n})d/2.$$

Now for each d, for sufficiently large n we have $\mu(G_d^{\le n}) \ge \delta$. As before, given r let $d = 2r/\delta$; then for large enough n we have $K(\mu\restriction_n) \le n + K(n) + c - r$. ◀

It would be interesting to know whether the above-mentioned coincidences of randomness notions for bit sequences lift to measures; for instance, do the conditions in the theorem above actually imply that the measure is ML-a.c. relative to the halting problem $\emptyset'$?

3.2 Both implications of the Levin-Schnorr Theorem fail for measures

We will show that both implications of the analog of the Levin-Schnorr Theorem 2 fail for measures. One implication would say that a Martin-Löf a.c. measure cannot have initial segment complexity in the sense of K growing slower than $n - O(1)$. This can be disproved by a simple example of a measure with countable support. On the other hand, by Proposition 25 below, we have $\lim_n K(\mu\restriction_n)/n = 1$ for each Martin-Löf a.c. measure μ, which provides a lower bound on the growth.

▶ **Example 12.** Let $\theta \in (0,1)$. There is a Martin-Löf a.c. measure μ such that $K(\mu\restriction_n) \le^+ n - n^\theta$.

Proof. We let $\mu = \sum c_k \delta_{Z_k}$ where Z_k is Martin-Löf random and $0^{n_k} \prec Z_k$ for a sequence $\langle c_k \rangle$ of reals in $[0,1]$ that add up to 1, and a sufficiently fast growing computable sequence $\langle n_k \rangle$ to be determined below. Then μ is Martin-Löf a.c. by Fact 7.

For n such that $n_k \le n < n_{k+1}$ we have

$$\begin{aligned} K(\mu\restriction_n) &\le^+ (\sum_{l=0}^{k} c_l)\cdot(n+2\log n) + (\sum_{l=k+1}^{\infty} c_l)\cdot 2\log n \\ &\le^+ (1-c_{k+1})n + 2\log n. \end{aligned}$$

Hence, to achieve $K(\mu\restriction_n) \le^+ n - n^\theta$ it suffices to ensure that $c_{k+1}n_k \ge n_{k+1}^\theta + 2\log n_{k+1}$ for almost all k. For instance, we can let $c_k = \frac{1}{(k+1)(k+2)}$ and $n_k = 2^{k+4}$. ◀

To disprove the converse implication, we need to provide a measure μ such that $K(\mu\restriction_n) \ge^+ n$ yet μ is not Martin-Löf a.c. This will be immediate from the following fact on the growth of initial segment complexity for certain bit sequences.

▶ **Theorem 13.** *There are a Martin-Löf random bit sequence X and a non-Martin-Löf random bit sequence Y such that, for all n, $K(X\restriction_n) + K(Y\restriction_n) \ge^+ 2n$.*

Proof. Let X be a low Martin-Löf random set (i.e., $X' \equiv_T \emptyset'$). We claim that there is a strictly increasing function f such that the complement of the range of f is a recursively enumerable set E, and $K(X\restriction_m) \ge m + 3n$ for all $m \ge f(n)$. To see this, recall that $\lim_n K(X\restriction_n) - n = \infty$. Since X is low there is a computable function p such that for all n, $\lim_s p(n,s)$ is the maximal m such that $K(X\restriction_m) \le m + 3n$.

Define $f(n,s)$ for $n \le s$ as follows. $f(n,0) = n$; for $s > 0$ let n be least such that $p(n,s) \ge f(n,s-1)$ or $n = s$. If $m \ge n$ and $n < s$ then let $f(m,s) = s + m - n$ else let $f(m,s) = f(m,s-1)$.

Note that for each n there are only finitely many $s > 0$ with $f(n,s) \ne f(n,s-1)$ and that almost all s satisfy $f(n,s) > p(n,s)$, as otherwise $f(n,s)$ would be modified either at n or some smaller value. Furthermore, $f(n,s) \ne f(n,s-1)$ can only happen if there is an $m \le n$ with $f(m,s-1) \le p(m,s)$ and that happens only finitely often, as all the $p(m,s)$ converge to a fixed value and every change of an $f(m,s)$ at some time s leads to a value above s. Furthermore, once an element is outside the range of f, it will never return, and so the complement of the range of f is recursively enumerable. So $f(n) = \lim_s f(n,s)$ is a function as required, which verifies the claim. (The complement E of the range of f is called a Dekker deficiency set in the literature [21].)

Now let $g(n) = \max\{m : f(m) \le n\}$ (with the convention that $\max(\emptyset) = 0$). Since g is unbounded, by a result of Miller and Yu [15, Cor. 3.2] there is a Martin-Löf random Z such that there exist infinitely many n with $K(Z\restriction_n) \le n + g(n)/2$; note that the result of Miller and Yu does not make any effectivity requirements on g. Let

$$Y = \{n + g(n) : n \in Z\}.$$

Note that $K(Z\restriction_n) \le K(Y\restriction_n) + g(n) + K(g(n))$, as one can enumerate the set E until there are, up to n, only $g(n)$ many places not enumerated and then one can reconstruct $Z\restriction_n$ from $Y\restriction_n$ and $g(n)$ and the last $g(n)$ bits of $Z\restriction_n$. As Z is Martin-Löf random, $K(Z\restriction_n) \ge^+ n$, so

$$K(Y\restriction_n) \ge^+ n - g(n) - K(g(n)) \ge^+ n - 2g(n).$$

The definitions of X, f, g give $K(X\restriction_n) \ge n + 3g(n)$. This shows that $K(X\restriction_n) + K(Y\restriction_n) \ge 2n$ for almost all n.

However, the set Y is not Martin-Löf random, as there are infinitely many n such that $K(Z\restriction_n) \le^+ n + g(n)/2$. Now $Y\restriction_{n+g(n)}$ can be computed from $Z\restriction_n$ and $g(n)$, as one needs only to enumerate E until the $g(n)$ nonelements of E below n are found and they allow to see where the zeroes have to be inserted into the string $Z\restriction_n$ in order to obtain $Y\restriction_{n+g(n)}$. Note furthermore, that $K(g(n)) \le g(n)/4$ for almost all n and thus $K(Y\restriction_{n+g(n)}) \le^+ n + 3/4 \cdot g(n)$ for infinitely many n, so Y cannot be Martin-Löf random. ◀

3.3 Failing a restricted type of test implies non-complex initial segments

We say that a Solovay test $\langle S_r \rangle_{r\in\mathbb{N}}$ is *strong* if each S_r is clopen and given by a strong index for a finite set of strings X_r such that $[X_r]^{\prec} = S_r$. For bit sequences, this means no restriction: any Solovay test can be replaced by a strong Solovay test listing the strings making up the Σ^0_1 sets one-by-one. (We conjecture that this equivalence of test notions no longer holds for measures.) The following is a weak version for measures of one implication of the Miller-Yu theorem [14, Thm. 7.1.]. We use elements of the proof in Bienvenu et al. [1]. We note that the result is a variation on (the contrapositive of) Theorem 10(a), proving a stronger conclusion from a stronger hypothesis.

▶ **Proposition 14.** *Suppose that μ fails a strong Solovay test $\langle S_r \rangle_{r\in\mathbb{N}}$ at level δ, namely $\exists^\infty r\,[\mu(S_r) \ge \delta]$. Then there is a computable function f such that $\sum_n 2^{-f(n)} < \infty$ and*

$$\exists^\infty n\,[C(\mu\restriction_n \mid n) \le^+ n - \delta f(n)].$$

Proof. Let $\langle X_r \rangle$ be as in the definition of a strong Solovay test. We may assume that $\sum \lambda S_r \le 1/2$, and all strings in X_r have the same length n_r. Let f be computable such that

$$2^{-f(n_r)} \ge \lambda(S_r) > 2^{-f(n_r)-1}$$

and $f(m) = 2^{-m}$ for m not of the form n_r. There is a constant d such that each bit string x in the set X_r satisfies (where $n = n_r$)

$$C(x \mid n) \le n - f(n) + d =: g(r). \tag{1}$$

For, r can be computed from $n = n_r$, and each string $x \in X_r$ is determined by r and its position $i < 2^n\lambda(X_r)$ in the lexicographical listing of X_r. We can determine i by $\log(2^n\lambda(X_r)) \le n - f(n) + d$ bits for some fixed d. In fact we may assume the description has exactly that many bits. Thus, there is a Turing machine L with two inputs such that for each $\sigma \in X_r$, we have $L(v_\sigma; n) = \sigma$ for some bit string v_σ of length $g(r)$.

Let c be a constant such that $C(x) \le |x| + c$ for each string x. Now suppose that $\mu(S_r) \ge \delta$. Then for $n = n_r$,

$$\begin{aligned} C(\mu\restriction_n \mid n) \quad &\le^+ \quad \sum_{\sigma\in X_r} (n - f(n))\mu[\sigma] + \sum_{\sigma\notin X_r} (n + c)\mu[\sigma] \\ &\le^+ \quad n - f(n) \sum_{\sigma\in X_r} \mu[\sigma] \\ &\le^+ \quad n - \delta f(n). \end{aligned}$$

Since there are infinitely many such r by hypothesis, this completes the proof. ◀

4 K-triviality for measures

▶ **Definition 15.** *A measure μ is called K-trivial if $K(\mu\restriction_n) \le^+ K(n)$ for each n.*

For Dirac measures δ_A this is the same as saying that A is K-trivial in the usual sense. More generally, any finite convex combination of such Dirac measures is K-trivial.

▶ **Proposition 16.** *If μ is K-trivial, then μ is supported by its set of atoms.*
In fact the weaker hypothesis that $\exists p \exists^\infty n\, [K(\mu\restriction_n) \le K(n) + p]$ suffices.

Proof. For a set $R \subseteq \{0,1\}^*$, by $[R]^\prec$ one denotes the open set $\{Z\colon\ \exists n\, Z\restriction_n \in R\}$.

Assume for a contradiction that μ gives a measure greater than $\epsilon > 0$ to the set of its non-atoms. Note that there is a constant b such that $K(x) \ge K(|x|) - b$ for each x. Fix c arbitrary with the goal of showing that $K(\mu\restriction_n) \ge K(n) - b + \epsilon c/2$ for large enough n.

There is d (in fact $d = O(2^c)$) such that for each n there are at most d strings x of length n with $K(x) \le K(n) + c$ (see e.g. [17, 2.2.26]). Let $S_n = \{x\colon\ |x| = n \wedge \mu[x] \le \epsilon/2d\}$. By hypothesis we have $\mu[S_n]^\prec \ge \epsilon$ for large enough n. Therefore by choice of d we have

$$\mu[S_n \cap \{x\colon\ K(x) > K(|x|) + c\}]^\prec \ge \epsilon/2.$$

Now we can give a lower bound for the μ-average of $K(x)$ over all strings x of length n:

$$\sum_{|x|=n} K(x)\mu[x] \ge (1-\epsilon/2)(K(n)-b) + (\epsilon/2)(K(n)+c) \ge K(n) - b + \epsilon c/2,$$

as required. Notice that we have only used the weaker hypothesis. ◀

Thus, if μ is K-trivial for constant p, then μ has the form $\sum_{r<N} \alpha_r \delta_{A_r}$ where $N \le \infty$ and each α_r is positive and $\sum_{r<N} \alpha_r = 1$.

▶ **Fact 17.** *Each A_r is K-trivial for constant $(p+c)/\alpha_r$, for some fixed c.*

Proof. Let c be a constant such that $K(x) \ge K(|x|) - c$ for each string x. We have

$$K(n) + p \ge K(\mu\restriction_n) = \sum_s \alpha_s K(A_s\restriction_n) \ge \alpha_r K(A_r\restriction_n) + \sum_{s \ne r} \alpha_s (K(n) - c).$$

Therefore $\alpha_r K(n) + p + c \ge \alpha_r K(A\restriction_n)$, as required. ◀

It would be interesting to characterise the countable convex combinations of K-trivials that yield K-trivial measures. The following is easily checked.

▶ **Fact 18.** *Suppose that A_r is K-trivial with constant b_r, and $\sum_r \alpha_r b_r \le c < \infty$ where each α_r is positive and $\sum \alpha_r = 1$. Then $\mu = \sum_r \alpha_r \delta_{A_r}$ is K-trivial with constant c.*

For instance, we can build a computable K-trivial measure with infinitely many atoms as follows. Let $A_r = 0^{r+1}1^\infty$, so that $K(A_r\restriction_n) \le^+ K(n) + 2\log r$. Let $\mu = \sum 2^{-r+1} A_r$. By the above fact μ is K-trivial. If we vary the construction by letting $A_r = 0^{r+1}1B$ where B is K-trivial but non-recursive, we obtain a K-trivial measure μ with infinitely many atoms, and none of them recursive.

On the other hand, the following example shows that not every infinite convex combination $\mu = \sum_k \alpha_k \delta_{A_k}$ of K-trivial Dirac measures for constant b_k yields a K-trivial measure, even if $\alpha_k b_k$ is bounded. Let $A_k = \{\ell : \ell \in \Omega \wedge \ell < k\}$, and $\alpha_k = (k+1)^{-1/2} - (k+2)^{-1/2}$. All sets A_k are finite and thus K-trivial for constant $2k + O(1)$. Furthermore, the sum of all α_k is 1 and $\alpha_k = O(1/k)$. We have

$$K(\mu\restriction_n) = \sum_{|x|=n} K(x)\mu(x) \ge (\sum_{m \ge n} \alpha_m) \cdot K(\Omega \restriction n) \ge (n+2)^{-1/2} \cdot (n+2) = \sqrt{n+2}$$

for almost all n, and thus the average grows faster than $K(n) + c$. So the measure is not K-trivial.

In a sense, an atomless measure can come arbitrarily close to being K-trivial.

▶ **Proposition 19.** *For each nondecreasing unbounded function f which is computably approximable from above there is a non-atomic measure μ such that $K(\mu\upharpoonright_n) \leq^+ K(n) + f(n)$.*

Proof. There is a recursively enumerable set A such that, for all n, $A \cap \{0, \ldots, n\}$ has up to a constant $f(n)/2$ non-elements. One lets μ be the measure such that $\mu(x) = 2^{-m}$ in the case that all ones in x are not in A and $\mu(x) = 0$ otherwise, where m is the number of non-elements of A below $|x|$. One can see that when $\mu(x) = 2^{-m}$ then x can be computed from $|x|$ and the string $b_0 b_1 \ldots b_{m-1}$ which describes the bits at the non-elements of A. Thus

$$K(x) \leq^+ K(|x|) + K(b_0 b_1 \ldots b_{m-1}) \leq^+ K(|x|) + 2m.$$

It follows that $K(\mu\upharpoonright_n) \leq^+ K(n) + f(n)$, as the μ-average of strings $x \in \{0,1\}^n$ with $K(x) \leq^+ K(n) + f(n)$ is at most $K(n) + f(n)$ plus a constant. ◀

5 Full Martin-Löf randomness of measures

Let $\mathcal{M}(\{0,1\}^{\mathbb{N}})$ be the space of probability measures on Cantor space (which is canonically a compact topological space). A probability measure $\mathbb{P}$ on this space has been introduced implicitly in Mauldin and Monticino [12]. Culver's thesis [4] shows that this measure is computable. So the framework of [8] yields a notion of Martin-Löf randomness for points in the space $\mathcal{M}(\{0,1\}^{\mathbb{N}})$.

To define $\mathbb{P}$, first let $\mathcal{R}$ be the closed set of representations of probability measures; namely, $\mathcal{R}$ consists of the functions $X\colon \{0,1\}^* \to [0,1]$ such that $X_\emptyset = 1$ and $X_\sigma = X_{\sigma 0} + X_{\sigma 1}$ for each string σ. $\mathbb{P}$ is the unique measure on $\mathcal{R}$ such that for each string σ and $r, s \in [0,1]$, we have $\mathbb{P}(X_{\sigma 0} \leq r \mid X_\sigma = s) = \min(1, r/s)$. Intuitively, we choose $X_{\sigma 0}$ at random w.r.t. the uniformly distribution on the interval $[0, X_\sigma]$, and the choices made at different strings are independent.

▶ **Proposition 20.** *Every probability measure μ that is Martin-Löf random wrt to $\mathbb{P}$ is Martin-Löf absolutely continuous.*

For the duration of this proof let μ range over $\mathcal{M}(\{0,1\}^{\mathbb{N}})$. For an open set $G \subseteq \{0,1\}^{\mathbb{N}}$, let

$$r_G = \int \mu(G) d\mathbb{P}(\mu).$$

Our proof of Prop. 20 is based on two facts.

▶ **Fact 21.** $r_G = \lambda(G)$.

Proof. Clearly, for each n we have

$$\sum_{|\sigma|=n} r_{[\sigma]} = \int \sum_{|\sigma|=n} \mu([\sigma]) d\mathbb{P}(\mu) = 1.$$

Furthermore, $r_\sigma = r_\eta$ whenever $|\sigma| = |\eta| = n$ because there is a $\mathbb{P}$-preserving transformation T of $\mathcal{M}(\{0,1\}^{\mathbb{N}})$ such that $\mu([\sigma]) = T(\mu)([\eta])$. Therefore $r_{[\sigma]} = 2^{-|\sigma|}$.

If σ, η are incompatible then $r_{[\sigma]\cup[\eta]} = r_{[\sigma]} + r_{[\eta]}$. Now it suffices to write $G = \bigcup_i [\sigma_i]$ where the strings σ_i are incompatible, so that $\lambda G = \sum_i 2^{-|\sigma_i|}$. ◀

▶ **Fact 22.** *Let $\mu \in \mathcal{M}(\{0,1\}^{\mathbb{N}})$ and let $\langle G_m \rangle_{m \in \mathbb{N}}$ be a ML-test such that there is $\delta \in \mathbb{Q}^+$ with $\forall m\, \mu(G_m) > \delta$. Then μ is not ML-random w.r.t. $\mathbb{P}$.*

Proof. Observe that by the foregoing fact

$$\delta \cdot \mathbb{P}(\{\mu\colon \mu(G_m) \geq \delta\}) \leq \int \mu(G_m) d\mathbb{P}(\mu) = \lambda(G_m) \leq 2^{-m}.$$

Let $\mathcal{G}_m = \{\mu\colon \mu(G_m) > \delta\}$ which is uniformly effectively open in the space of measures $\mathcal{M}(\{0,1\}^{\mathbb{N}})$. Fix k such that $2^{-k} \leq \delta$. By the inequality above, we have $\mathbb{P}(\mathcal{G}_m) \leq 2^{-m}/\delta \leq 2^{-m+k}$. Hence $\langle \mathcal{G}_{m+k} \rangle_{m \in \mathbb{N}}$ is a ML-test w.r.t. $\mathbb{P}$ that succeeds on μ. ◀

This argument also works for randomness notions stronger than Martin-Löf's. For instance, if there is a weak-2 test $\langle G_m \rangle_{m \in \mathbb{N}}$ such that $\mu G_m > \delta$ for each m, then μ is not weakly 2-random with respect to $\mathbb{P}$. The converse of Prop. 20 fails. Culver [4] shows that each measure μ that is Martin-Löf random w.r.t. $\mathbb{P}$ is non-atomic. So a measure δ_Z for a Martin-Löf random bit sequences Z is Martin-Löf a.c. but not Martin-Löf random with respect to $\mathbb{P}$.

6 Being ML-a.c. relative to computable ergodic measures

We review some notions from the field of symbolic dynamics, a mathematical area closely related to Shannon information theory. We will consider effective "almost-everywhere theorems" related to that area in the framework of randomness for measures.

It can be useful to admit alphabets other than the binary one. Let $\mathbb{A}^\infty$ denote the topological space of one-sided infinite sequences of symbols in an alphabet $\mathbb{A}$. Randomness notions etc. carry over from the case of $\mathbb{A} = \{0,1\}$. A dynamics on $\mathbb{A}^\infty$ is given by the shift operator T, which erases the first symbol of a sequence. A measure ρ on $\mathbb{A}^\infty$ is called *shift invariant* if $\rho(G) = \rho(T^{-1}(G))$ for each open (and hence each measurable) set G. The *empirical entropy* of a measure ρ along $Z \in \mathbb{A}^\infty$ is given by the sequence of random variables

$$h_n^\rho(Z) = -\frac{1}{n} \log_{|\mathbb{A}|} \rho[Z \upharpoonright_n].$$

A shift invariant measure ρ on $\mathbb{A}^\infty$ is called *ergodic* if every ρ integrable function f with $f \circ T = f$ is constant ρ-almost surely. The following equivalent condition can be easier to check: for any strings $u, v \in \mathbb{A}^*$,

$$\lim_N \frac{1}{N} \sum_{k=0}^{n-1} \rho([u] \cap T^{-k}[v]) = \rho[u]\rho[v].$$

For ergodic ρ, the entropy $H(\rho)$ is defined as $\lim_n H_n(\rho)$, where

$$H_n(\rho) = -\frac{1}{n} \sum_{|w|=n} \rho[w] \log \rho[w].$$

Thus, $H_n(\rho) = \mathbb{E}_\rho h_n^\rho$ is the expected value with respect to ρ. One notes that $H_{n+1}(\rho) \leq H_n(\rho) \leq 1$ so that the limit $H(\rho)$ exists.

A well-known result from the 1950s due to Shannon, McMillan and Breiman states that for an ergodic measure ρ, for ρ-a.e. Z the empirical entropy along Z converges to the entropy of the measure. See e.g. [24], but note that the result is called the Entropy Theorem there.

▶ **Theorem 23** (SMB theorem, e.g. [24]). *Let ρ be an ergodic measure on the space $\mathbb{A}^\infty$. For ρ-almost every Z we have $\lim_n h_n^\rho(Z) = H(\rho)$.*

A measure ρ on $\mathbb{A}^\infty$ is called *computable* if each real $\rho[x]$ is computable, uniformly in $x \in \mathbb{A}^*$. For such a measure we can define Martin-Löf tests and Martin-Löf randomness with respect to ρ (called ρ-ML randomness for short) as above. Recall from Fact 4 that a measure μ is *Martin-Löf a.c. with respect to* ρ iff $\mu(\mathcal{C}) = 0$ where $\mathcal{C}$ is the class of sequences in $\mathbb{A}^\infty$ that are not ML-random with respect to ρ.

If a computable measure ρ is shift invariant, then $\lim_n h^\rho_n(Z)$ exists for each ρ-ML-random Z by a result of Hochman [6]. Hoyrup [7, Thm. 1.2] gave an alternative proof for ergodic ρ, and also showed that in that case we have $\lim_n h^\rho_n(Z) = H(\rho)$ for each ρ-ML random Z. We extend this result to measures μ that are Martin-Löf a.c. with respect to ρ, under the additional hypothesis that the h^ρ_n are uniformly bounded. This holds e.g. for Bernoulli measures and the measures given by a Markov process. On the other hand, using a renewal process it is not hard to construct an ergodic computable measure ρ over the binary alphabet where this hypothesis fails. For instance, take a computable sequence of rationals $\langle \alpha_k \rangle$ with sum 1 which decreases quickly enough such that $\lim_k -\frac{1}{k+2} \log_2 \alpha_k = \infty$, and let ρ be a shift invariant measure such that $\rho(10^k 1 \prec Z \mid Z_0 = 1) = \alpha_k$ for each $k \in \mathbb{N}$. This method yields an example showing that the boundedness hypothesis in the proposition below is necessary. See the Logic Blog 2020 posted from Nies' website.

▶ **Proposition 24.** *Let ρ be a computable ergodic measure on the space $\mathbb{A}^\infty$ such that for some constant D, each h^ρ_n is bounded above by D. Suppose the measure μ is Martin-Löf a.c. with respect to ρ. Then $\lim_n E_\mu h^\rho_n = H(\rho)$.*

Proof. By Hoyrup's result, $\lim_n h^\rho_n(Z) = H(\rho)$ for each ρ-ML random Z. Since the sequences that are not ML-random w.r.t. ρ form a null set w.r.t. μ, we infer that $\lim_n h^\rho_n(Z) = H(\rho)$ for μ-a.e. Z. The exception set V is measurable. Let $\widetilde{h}^\rho_n$ be the function obtained by changing the value of h^ρ_n to 0 on this set. Then $\widetilde{h}^\rho_n(Z) \to H(\rho) 1_{(\mathbb{A}^\infty \setminus V)(Z)}$ for each Z. The Dominated Convergence Theorem now shows that $\lim_n \mathbb{E}_\mu h^\rho_n = H(\rho)$, as required. ◀

The next observation shows that the asymptotic initial segment complexity of a ML-a.c. measure relative to ρ obeys some lower bound. Note that $H(\lambda) = 1$. So for $\rho = \lambda$, this shows that in Example 12 we cannot subtract, say, $n/4$ instead of n^θ.

▶ **Proposition 25.** *Let ρ be a computable ergodic measure, and suppose μ is a Martin-Löf a.c. measure with respect to ρ. Then*

$$\lim_n \frac{1}{n} K(\mu\restriction_n) = \lim_n \frac{1}{n} C(\mu\restriction_n) = H(\rho).$$

Proof. We can use K and C interchangeably because $C(x) \leq^+ K(x) \leq^+ C(x) + K(C(x))$ [17, 2.4.1]. We choose K. Let $k_n(Z) = K(Z\restriction_n)/n$. The argument is very similar to the one in the theorem above, replacing the functions h_n by k_n. Note that k_n is bounded above by a constant because $K(x) \leq^+ |x| + 2\log|x|$. Hoyrup's result [7, Thm. 1.2] states that $\lim_n k_n(Z) = H(\rho)$ for each ρ-ML random Z. Now we can apply the Dominated Convergence Theorem as in the proof of the foregoing proposition. ◀

A further interesting direction to tackle in the measure case would be the effective Birkhoff's ergodic theorem. This says that for ergodic computable ρ, if $f\{0,1\}^\mathbb{N} \to \mathbb{R}$ is ρ-integrable and lower semicomputable and Z is ρ-ML-random, then the limit of the usual ergodic averages $A_n f(Z) = \frac{1}{n} \sum_{k<n} (f \circ T^k)(Z)$ equals $\int f d\rho$. (For background see e.g. [16] which contains references to original work.) If the $A_n f$ are bounded then an argument similar to the one above shows that $\lim_n \int A_n f d\mu = \int f d\rho$ for any $\mu \ll_{ML} \rho$, but without this additional hypothesis the question remains open.

References

1 Laurent Bienvenu, Wolfgang Merkle, and Alexander Shen. A simple proof of the Miller-Yu Theorem. *Fundamenta Informaticae*, 83(1-2):21–24, 2008.

2 Cristian Calude. *Information and randomness*. Monographs in Theoretical Computer Science. An EATCS Series. Springer-Verlag, Berlin, 1994. With forewords by Gregory J. Chaitin and Arto Salomaa.

3 Gregory J. Chaitin. A theory of program size formally identical to information theory. *J. Assoc. Comput. Mach.*, 22:329–340, 1975.

4 Quinn Culver. *Topics in Algorithmic Randomness and Effective Probability*. PhD thesis, PhD thesis, University of Notre Dame, 2015.

5 Rodney G. Downey and Denis Hirschfeldt. *Algorithmic randomness and complexity*. Springer-Verlag, Berlin, 2010. 855 pages.

6 Michael Hochman. Upcrossing inequalities for stationary sequences and applications. *Annals of Probability*, 37(6):2135–2149, 2009.

7 Mathieu Hoyrup. The dimension of ergodic random sequences. In Christoph Dürr and Thomas Wilke, editors, *STACS*, pages 567–576, 2012.

8 Mathieu Hoyrup and Cristóbal Rojas. Computability of probability measures and Martin-Löf randomness over metric spaces. *Information and Computation*, 207(7):830–847, 2009.

9 Leonid A. Levin. The concept of a random sequence. *Dokl. Akad. Nauk SSSR*, 212:548–550, 1973.

10 Ming Li and Paul Vitányi. *An introduction to Kolmogorov complexity and its applications*. Graduate Texts in Computer Science. Springer-Verlag, New York, second edition, 1997.

11 Per Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619, 1966.

12 Daniel Mauldin and Michael Monticino. Randomly generated distributions. *Israel Journal of Mathematics*, 91(1-3):215–237, 1995.

13 Joseph S. Miller. The K-degrees, low for K-degrees, and weakly low for K sets. *Notre Dame J. Form. Log.*, 50(4):381–391, 2009. doi:10.1215/00294527-2009-017.

14 Joseph S. Miller and Liang Yu. On initial segment complexity and degrees of randomness. *Trans. Amer. Math. Soc.*, 360:3193–3210, 2008.

15 Joseph S. Miller and Liang Yu. Oscillation in the initial segment complexity of random reals. *Advances in Mathematics*, 226(6):4816–4840, 2011.

16 Kenshi Miyabe, André Nies, and Jing Zhang. Using almost-everywhere theorems from analysis to study randomness. *Bull. Symb. Logic*, 22:305–331, 2016. arXiv:1411.0732.

17 André Nies. *Computability and Randomness*, volume 51 of *Oxford Logic Guides*. Oxford University Press, Oxford, 2009. 444 pages. Paperback version 2011. doi:10.1093/acprof:oso/9780199230761.001.0001.

18 André Nies and Volkher Scholz. Martin-Löf random quantum states. *Journal of Mathematical Physics*, 60(9):092201, 2019. Available at doi.org/10.1063/1.5094660.

19 André Nies, Frank Stephan, and Sebastiaan Terwijn. Randomness, relativization and Turing degrees. *J. Symbolic Logic*, 70(2):515–535, 2005.

20 Editor André Nies. Logic Blog 2017. Available at arXiv:1804.05331, 2017.

21 Piergiorgio Odifreddi. *Classical Recursion Theory (Volume I)*. North–Holland Publishing Co., Amsterdam, 1989.

22 Piergiorgio Odifreddi. *Classical recursion theory: The theory of functions and sets of natural numbers*, volume 125. Elsevier, 1992.

23 Claus-Peter Schnorr. Process complexity and effective random tests. *J. Comput. System Sci.*, 7:376–388, 1973. Fourth Annual ACM Symposium on the Theory of Computing (Denver, Colo., 1972).

24 Paul C. Shields. *The Ergodic Theory of Discrete Sample Paths*. Graduate Studies in Mathematics 13. American Mathematical Society, 1996.

25 Robert I. Soare. *Recursively Enumerable Sets and Degrees*. Perspectives in Mathematical Logic, Omega Series. Springer–Verlag, Heidelberg, 1987.

26 Frank Stephan. Recursion theory. Lecture notes, School of Computing, 2012.